Fisheries and Oceans Canada | Pêches et Océans Canada

Canadian Stock Assessment Secretariat
Research Document 99/110

Not to be cited without
permission of the authors[1]

Secrétariat canadien pour l'évaluation des stocks
Document de recherche 99/110

Ne pas citer sans
autorisation des auteurs[1]

Inseason forecast model for 1SW salmon returning to the Saint John River, New Brunswick, in 1999

by

C.J. Harvie and T.L. Marshall

Science Branch
Department of Fisheries and Oceans
P.O. Box 1006
Dartmouth, N.S.
B2Y 4A2

[1] This series documents the scientific basis for the evaluation of fisheries resources in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Research documents are produced in the official language in which they are provided to the Secretariat.

[1] La présente série documente les bases scientifiques des évaluations des ressources halieutiques du Canada. Elle traite des problèmes courants selon les échéanciers dictés. Les documents qu'elle contient ne doivent pas être considérés comme des énoncés définitifs sur les sujets traités, mais plutôt comme des rapports d'étape sur les études en cours.

Les documents de recherche sont publiés dans la langue officielle utilisée dans le manuscrit envoyé au secrétariat.

ISSN 1480-4883
Ottawa, 1999
Canada

Abstract

Inseason forecasts of wild and hatchery 1SW Atlantic salmon returning to Mactaquac Dam on the Saint John River, New Brunswick, have been estimated using simple linear regression models. The response variable has been cumulative proportion of the returns to date and, for wild 1SW salmon, the predictor variable has been average daily mean discharge. Since 1995, discharge has become a less significant predictor in the wild 1SW model and the seasonal distribution of hatchery 1SW arrival has changed, frequently resulting in overestimation of returns. A new model, regressing end-of-season returns on returns-to-date, for both wild and hatchery components at three dates, was evaluated for forecast accuracy in a retrospective examination and found to more accurately forecast end-of-season returns.

Résumé

Les prévisions en cours de saison des remontées de saumons de l'Atlantique sauvages et d'élevage unibermarins au barrage Mactaquac sur la rivière Saint-Jean, au Nouveau-Brunswick, ont été estimées par modèles simples de régression linéaire. La variable de réponse a été la proportion cumulative des remontées actualisées et, pour les saumons sauvages unibermarins, la variable prédictive a été l'écoulement quotidien moyen. Depuis 1995, l'écoulement est devenu une variable prédictive moins significative pour le modèle des saumons sauvages unibermarins et la distribution saisonnière des arrivées des unibermarins d'élevage s'est modifiée donnant souvent lieu à une surestimation des retours. Un nouveau modèle, de régression des remontées en fin de saison sur les remontées actualisées, appliqué aux éléments sauvage et d'élevage à trois dates particulières, a été évalué en ce qui a trait à l'exactitude des prévisions par examen rétrospectif et cela a permis d'obtenir une prévision plus juste des remontées de fin de saison.

Background

Inseason forecasts of 1SW Atlantic salmon returning to Mactaquac Dam on the Saint John River have been estimated using a simple linear regression of cumulative proportion of the run returning to date on the average daily mean discharge over the period July 2 to one day before the forecast date (Marshall and Cameron MS 1995). The model is of the form:

$$\text{Arcsine cumulative proportion to date} = \beta_0 + \beta_1 * \text{Average daily mean discharge} + \text{error}$$

A forecast equation was estimated for wild and hatchery 1SW salmon separately, for each of three dates: July 15, 22, and 29. Forecast equations estimated for wild 1SW salmon before 1995 indicated significant regression coefficients ($p<0.05$) for average daily mean discharge in the model, for each forecast date. Equations forecasting hatchery 1SW salmon returns over the same period were found to be not significantly dependent on discharge, so the model took the form:

$$\text{Arcsine cumulative proportion to date} = \beta_0 + \text{error} \qquad (1)$$

That is, the historical (1972 to forecast year-1) average arcsine cumulative proportion to date was used to estimate the forecast of hatchery 1SW salmon returns.

Starting in 1995, it was noted that the seasonal distribution of hatchery 1SW returns had changed. In addition, the regression coefficient for average daily mean discharge in the wild 1SW salmon model was no longer significant ($p>0.05$) for the July 15 forecast date. A model similar to the hatchery model (1) was used for this date. In 1997, none of the three forecast dates produced equations with significant regression coefficients for the independent variable, for either wild or hatchery 1SW returns. Model (1) was used to make all forecasts. However, this model often overestimated returns, particularly at medium to high values of returns-to-date.

In 1998, new potential forecast models were explored to determine whether a model could be found which produced more accurate estimates of returns, particularly at medium to high values of returns-to-date. Regression equation coefficients for the models, for each of wild and hatchery 1SW returns at each forecast date, were estimated by an iteratively reweighted least squares (IRLS) robust regression procedure (Neter *et al.* 1996) to reduce the influence of outliers. This examination produced the model:

$$\text{End of season returns} = \beta_0 + \beta_1 * \text{Returns to date} + \text{error} \qquad (2)$$

The 1998 returns-to-date values were very low for wild returns and very high for hatchery returns. At these values, model (2) would overestimate wild returns and model (1) would overestimate hatchery returns, particularly on July 15 and 22. Therefore, 1998 forecasts were made using model (1) for wild returns and model (2) for hatchery returns.

With the addition of the 1998 data point to the data set, models (1) and (2) became more similar in their forecasts of wild 1SW returns. The two models were compared to determine the one model which would best fit the data by providing more accurate estimates of returns. The same was examined for hatchery 1SW returns. For each of models (1) and (2), evaluations of forecast accuracy were made by estimating, for each of the years 1972 to 1998, a forecast returns value and comparing this to the actual returns value. The sum, over all years, of the squared differences between the actual and forecast returns was calculated to give an indication of the overall accuracy of forecasting with the two models.

Results

Regression coefficients and statistics were estimated for each of wild and hatchery 1SW returns at each forecast date (Figures 1-3). The discharge variable did not add significantly to any of the equations. The regression equations are:

Wild 1SW returns forecast from July 15

Model (1): Arcsine cumulative proportion returns to July 15 = 32.682 (n=27)

Model (2): End of season returns = 1229.701 + 2.375 * Returns to July 15 (R^2_{adj}=0.646; p<0.001; n=27)

Wild 1SW returns forecast from July 22

Model (1): Arcsine cumulative proportion returns to July 22 = 45.452 (n=27)

Model (2): End of season returns = 526.736 + 1.646 * Returns to July 22 (R^2_{adj}=0.858; p<0.001; n=27)

Wild 1SW returns forecast from July 29

Model (1): Arcsine cumulative proportion returns to July 29 = 54.305 (n=27)

Model (2): End of season returns = 90.826 + 1.441 * Returns to July 29 (R^2_{adj}=0.954; p<0.001; n=27)

Hatchery 1SW returns forecast from July 15

Model (1): Arcsine cumulative proportion returns to July 15 = 24.816 (n=27)

Model (2): End of season returns = 1263.058 + 2.275 * Returns to July 15 (R^2_{adj}=0.577; p<0.001; n=27)

Hatchery 1SW returns forecast from July 22

Model (1): Arcsine cumulative proportion returns to July 22 = 35.826 (n=27)

Model (2): End of season returns = 886.891 + 1.703 * Returns to July 22 (R^2_{adj}=0.732; p<0.001; n=27)

Hatchery 1SW returns forecast from July 29

Model (1): Arcsine cumulative proportion returns to July 29 = 44.506 (n=27)

Model (2): End of season returns = 501.354 + 1.580 * Returns to July 29 (R^2_{adj}=0.845; p<0.001; n=27)

The sum of squared differences between forecast returns and actual returns, for each of models (1) and (2), clearly indicates that model (2) more accurately forecasts returns of both wild and hatchery 1SW salmon, when only one model is used (Tables 1-2). Therefore, model (2) will be used to estimate 1999 forecasts of wild and hatchery 1SW salmon returns. The sum of squared differences for the July 29 forecast of wild returns indicated very little difference between the two models. For consistency, model (2) was chosen.

References

Marshall, T.L. and J D. Cameron. MS 1995. Status of Atlantic salmon stocks of the Saint John River and southwest New Brunswick, 1994. DFO Atlantic Fisheries Res. Doc. 95/129.

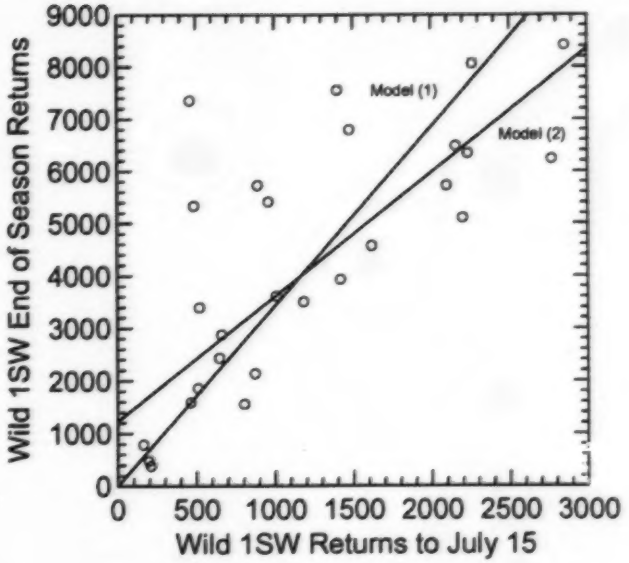
Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman. 1996. Applied Linear Statistical Models. Fourth Edition. IRWIN, Inc., USA.

Table 1. End of season returns and forecast returns of wild 1SW salmon returning to Mactaquac Dam on the Saint John River, 1972 to 1998, for each of three forecast dates and two models.

	Wild 1SW Returns									
	Actual End of Season Returns	Returns to Date			Forecast End of Season Returns					
					Model (1)			Model (2)		
Year		July 15	July 22	July 29	July 15	July 22	July 29	July 15	July 22	July 29
1972	784	164	275	340	562	541	515	1619	979	581
1973	1854	511	1008	1272	1753	1985	1929	2443	2186	1924
1974	3389	522	1555	2208	1790	3062	3348	2469	3086	3273
1975	5725	2097	3364	4006	7192	6623	6074	6210	6064	5863
1976	6797	1477	2973	4186	5066	5854	6347	4738	5420	6123
1977	3506	1183	2084	2530	4057	4103	3836	4039	3957	3737
1978	1584	464	850	993	1591	1674	1506	2332	1926	1522
1979	6234	2770	3825	4365	9500	7531	6618	7808	6823	6381
1980	7555	1401	2406	4076	4805	4737	6180	4557	4487	5964
1981	4571	1619	2608	3301	5553	5135	5005	5075	4820	4848
1982	3931	1419	2033	2601	4867	4003	3944	4600	3873	3839
1983	3613	1010	1802	2459	3464	3548	3728	3628	3493	3634
1984	7353	463	1521	2595	1588	2995	3934	2329	3030	3830
1985	5331	487	1157	2403	1670	2278	3643	2386	2431	3554
1986	6347	2233	3215	4519	7658	6330	6851	6533	5819	6603
1987	5106	2200	3068	3675	7545	6041	5572	6455	5577	5387
1988	8062	2264	4711	5689	7765	9276	8625	6607	8281	8289
1989	8417	2855	5321	6308	9792	10477	9564	8010	9285	9181
1990	6486	2157	3390	4059	7398	6675	6154	6353	6107	5940
1991	5415	960	1794	2608	3292	3532	3954	3510	3480	3849
1992	5729	893	2236	3545	3063	4403	5375	3351	4207	5199
1993	2873	661	1283	2038	2267	2526	3090	2800	2639	3028
1994	2133	873	1476	1670	2994	2906	2532	3303	2956	2497
1995	2429	648	1227	1721	2222	2416	2609	2769	2546	2571
1996	1552	804	1065	1240	2757	2097	1880	3139	2280	1878
1997	380	211	279	299	724	549	453	1731	986	522
1998	476	198	310	370	679	610	561	1700	1037	624
Sum of squared differences:					99,655,057	53,500,470	22,127,013	74,060,337	48,969,418	22,931,755

Table 2. End of season returns and forecast returns of hatchery 1SW salmon returning to Mactaquac Dam on the Saint John River, 1972 to 1998, for each of three forecast dates and two models.

	Hatchery 1SW Returns									
	Actual End of Season Returns	Returns to Date			Forecast End of Season Returns					
					Model (1)			Model (2)		
Year		July 15	July 22	July 29	July 15	July 22	July 29	July 15	July 22	July 29
1972	246	21	48	66	119	140	134	1311	969	606
1973	1760	375	755	1013	2129	2204	2062	2116	2173	2102
1974	3700	275	1058	1603	1561	3088	3262	1889	2689	3034
1975	5335	1380	2455	2970	7834	7166	6044	4403	5068	5193
1976	7694	415	1101	2402	2356	3214	4888	2207	2762	4296
1977	6201	1571	3021	3828	8918	8818	7790	4837	6032	6548
1978	2556	407	858	1115	2310	2504	2269	2189	2348	2263
1979	3521	1000	1448	1745	5677	4226	3551	3538	3353	3258
1980	9759	1290	2204	3889	7323	6433	7915	4198	4641	6645
1981	3782	1004	1709	2247	5700	4988	4573	3547	3798	4051
1982	2292	310	505	800	1760	1474	1628	1968	1747	1765
1983	1230	207	362	575	1175	1057	1170	1734	1503	1410
1984	1304	72	205	400	409	598	814	1427	1236	1133
1985	1746	68	207	462	386	604	940	1418	1239	1231
1986	699	136	225	364	772	657	741	1572	1270	1076
1987	2894	760	1164	1496	4314	3397	3045	2992	2869	2865
1988	1129	266	518	692	1510	1512	1408	1868	1769	1594
1989	1170	157	420	571	891	1226	1162	1620	1602	1403
1990	1421	215	436	566	1221	1273	1152	1752	1630	1395
1991	2160	74	207	411	420	604	836	1431	1239	1151
1992	1935	109	406	850	619	1185	1730	1511	1578	1844
1993	1034	106	306	619	602	893	1260	1504	1408	1479
1994	1180	260	567	725	1476	1655	1475	1855	1853	1647
1995	2541	321	737	1236	1822	2151	2515	1993	2142	2454
1996	4603	1854	2346	2681	10525	6848	5456	5481	4883	4736
1997	2689	1390	1949	2267	7891	5689	4614	4425	4207	4082
1998	4413	2346	3278	3696	13318	9568	7522	6600	6470	6340
Sum of squared differences:					213,373,017	90,789,026	32,844,281	80,862,484	62,343,238	30,448,033



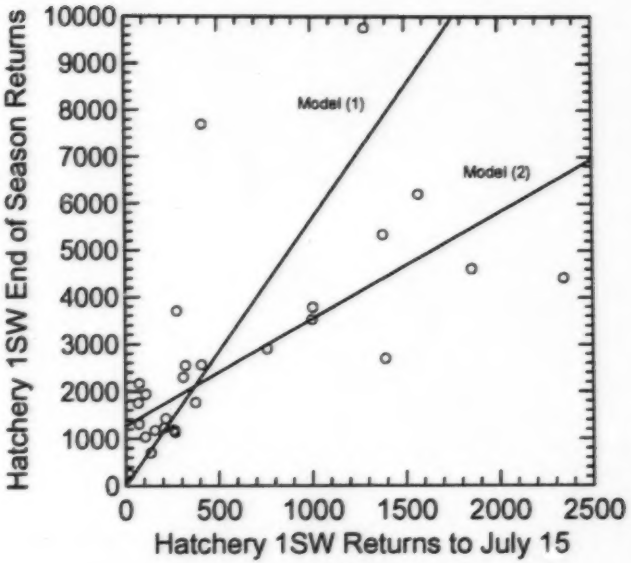


Figure 1. Plot of wild (left) and hatchery (right) 1SW end of season returns against returns to July 15, 1972 to 1998, on the Saint John River.

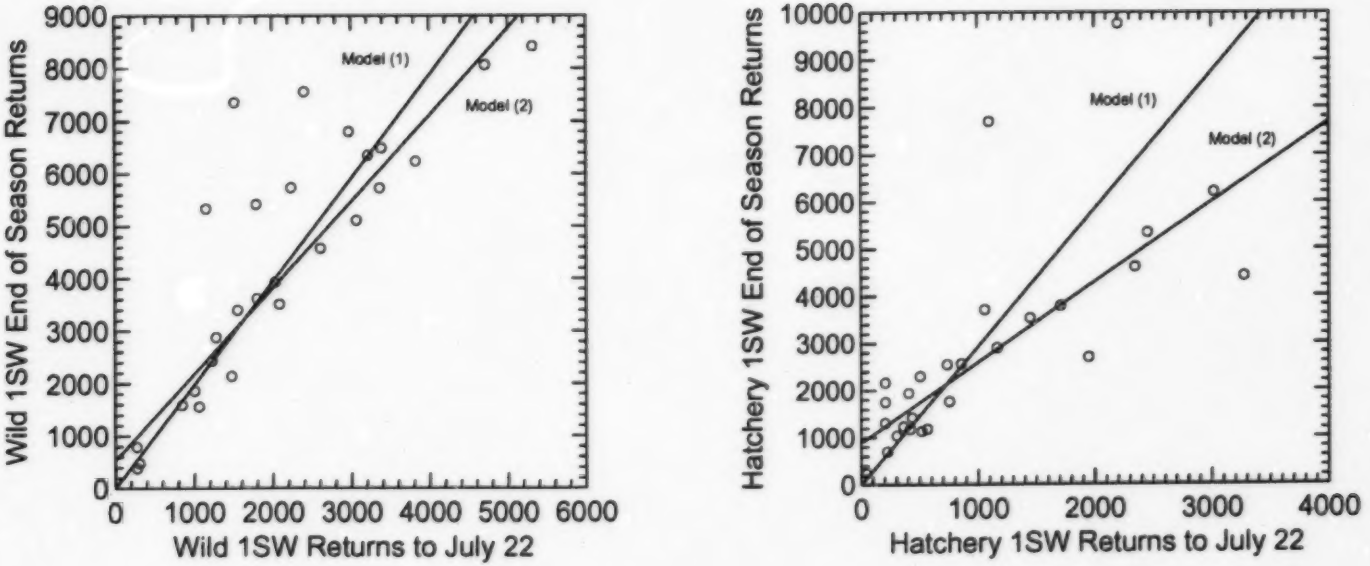


Figure 2. Plot of wild (left) and hatchery (right) 1SW end of season returns against returns to July 22, 1972 to 1998, on the Saint John River.

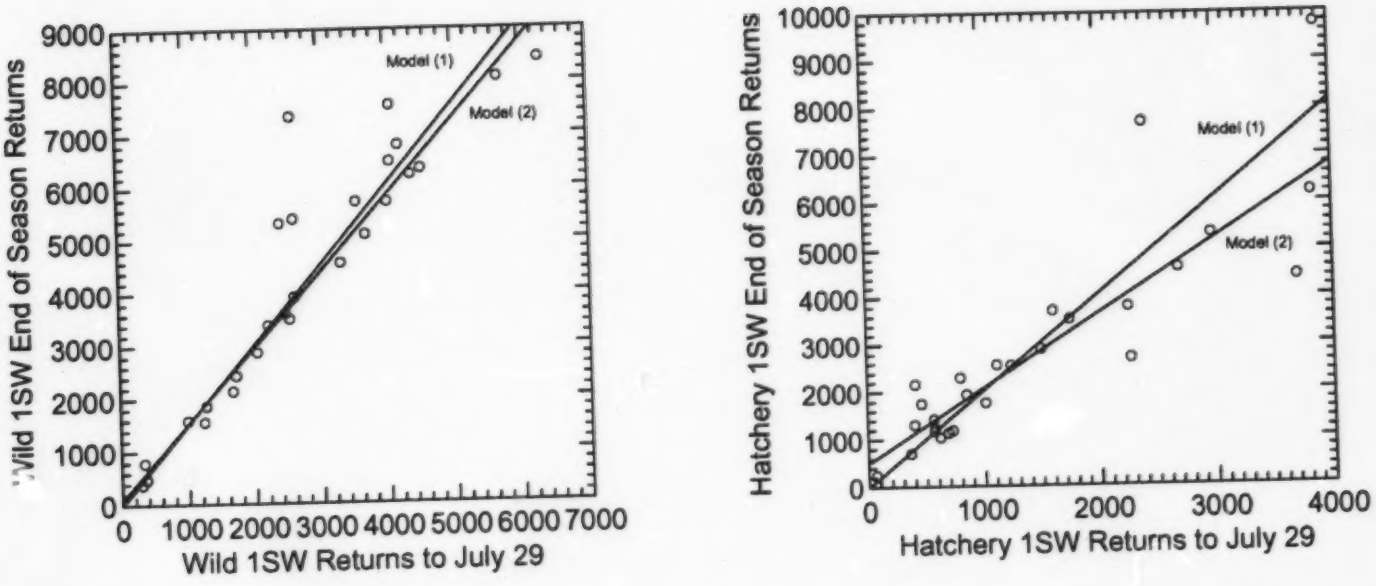


Figure 3. Plot of wild (left) and hatchery (right) 1SW end of season returns against returns to July 29, 1972 to 1998, on the Saint John River.